Int. J. Open Problems Compt. Math., Vol. 3, No. 4, December 2010
ISSN 1998-6262;
www.i-csrs.org

Learning-Based Document Image Super-Resolution with Directional Total Variation

Osama A. Omer

Department of Electrical Engineering,
South Valley University, Aswan, Egypt.
e-mail:omer.osama@gmail.com

Abstract

We propose a super-resolution algorithm based on local adaptation. In the proposed algorithm, the mapping function from the low-resolution images to high-resolution image is estimated by adaptation. Moreover, the property of the high-resolution image is learned and incorporated in a regularization -based restoration. The proposed regularization function is used as a general directional total variation with adaptive weights. The adaptive weights of the directional total variation are estimated based on the property of the partially reconstructed high-resolution image. The regularization function can be thought as a linear combination of smoothness in different directions. The convexity conditions as well as the convergence conditions are studied for the proposed algorithm. The proposed algorithm is tested for document images with various types of data, including plots, equations, draws and texts, existing in real documents.

Keywords: *Super-resolution, image fusion, restoration, directional total variation, regularization, document image.*

1 Introduction

The problem of image super-resolution (SR), where one obtains high-resolution (HR) image(s) from a set of low-resolution (LR) images, has many applications in the fields of image processing and computer vision. In particular, document analysis systems are becoming increasingly visible in daily life applications [1]. For instance, one may be interested in systems that process, store, understand document images. Resolution enhancement of LR document images is becoming an important pre-requisite for design and development of robust document analysis systems.

Large scale camera-based book scanners employed in digital libraries require resolution enhancement to obtain high optical character recognition (OCR) accuracies. The task of resolution enhancement is typically to increase spatial resolution which can assist the recognition in LR document images.

Methods of image SR can be divided into two classes. Namely, multi-images SR and single image SR. In the former class, the HR image is obtained by fusing information from multiple LR images to get a single image with higher resolution while in the later class, one extracts HR image details from a single LR image, which cannot be achieved by simple interpolation and sharpening.

Many approaches for image SR from multiple LR images have been presented in the last two decades [2] including recursive least squares (RLS) [3], iterative back-projection (IBP) [4, 5], maximum likelihood (ML) [6], maximum a posteriori (MAP) [7, 8], projection onto convex sets (POCS) [9, 10], nonuniform interpolation [11], FIR filter [12], machine learning [13, 14] and neural networks [15–17]. Among these algorithms, the two steps-based SR algorithms [18, 19] are reviewed and extended in this paper for its simplicity.

In [18, 19], the SR process has been performed by two steps, namely, the fusion step and restoration step. In [18], the LR images are fused by shift-and-add, where the up-sampled and warped images are averaged to get a blurred version of the HR image which is restored in the restoration step. In this algorithm, it is assumed that all the LR images contain HR information. In fact this assumption is not always correct due to the registration error. On the other hand, the median-shift-and-add (MSA) is used in [19] to fuse the LR images. Indeed, taking median of the up-sampled and warped images is more robust against the registration error, however, median is effective in case that less than 50% are error pixels which is not guaranteed in case of two sources of noise (the registration error and additive noise). Least mean square (LMS) estimator has been used in weighted image gradient fusion in [12]. However, this algorithm greatly depends on the desired response of the LMS estimator. In this algorithm, median values are used as the desired solution which can lead to local minima of the cost function.

In addition, the HR image is reconstructed by performing a restoration step to the fused images. Because of the ill-posedness of the restoration problem, typically the ill-posed problem is solved by adding constraint(s) to the solution by adding a regularization term to the cost function. In [12, 18], the Tikhonov regularization is used as a regularization function. In [19], the bilateral total variation (BTV) is used as a regularization function. The use of Tikhonov regularization implicitly assumes smoothness of the HR image in all directions which is not always correct. On the other hand, BTV regularization can overcome smoothness cross sharp edges by adjusting the bilateral filters' parameters. However, the main problem of BTV is the great dependency on the bilateral filters' parameters. Moreover, the weight for smoothness in each direction has a fixed value for constant bilateral parameters regardless to the image's property.

Hence, the motivation of this work is to develop an unsupervised learning-based SR algorithm that can overcome the problems of both fusion step and restoration step. To do that, the mapping function between the LR and HR images is estimated by learning in the fusion step which can overcome the problem of using fixed mapping function. Moreover, in the restoration step, rather than using Tikhonov or BTV as a regularization function, a general *directional total variation* (DTV) function with adaptive weights is newly introduced in a regularization-based restoration. The weights for different directions are adaptively estimated based on the partially estimated HR image. This weight adaptation can sense the image's local properties and then overcomes the problem of using fixed smoothness weight for different direction.

The remaining sections of this paper are organized as follows. Section 2 describes the mathematical model for the super-resolution problem. Section 3 briefly review a state-of-the-art two steps-based SR algorithm and the problems of this algorithm are highlighted. The proposed two steps-based SR algorithm is introduced in Section 5. Section 6 presents the simulation results on the fusion, restoration and SR of document image sequences. The conclusions are presented in Section 7. Finally, Section 8 suggests an extension to the proposed algorithm.

2 Problem Description

Assume that K LR images of the same scene in Lexicographical order denoted by $Y_k (1 \leq k \leq K)$, each containing M^2 pixels, are observed, and they are generated from the HR image denoted by X, containing L^2 pixels, where $L \geq M$. The observation of K LR images are modeled by the following degradation process:

$$Y_k = \mathbf{DHF}_k X + V_k. \tag{1}$$

where $\mathbf{F}_k$, $\mathbf{H}$ and $\mathbf{D}$ are the motion operator of the k^{th} image, the blurring operator (due to camera), and the down-sampling operator respectively, X is the unknown HR image, Y_k is the k^{th} observed LR image, and V_k is an additive random noise for the k^{th} image. Throughout the paper, we assume that $\mathbf{D}$ and $\mathbf{H}$ are known and the additive noise is Gaussian with zero mean. Therefore the problem here in this paper is to find the original image X.

3 Two Steps-Based Super-Resolution

Traditionally, the SR problem is solved by using regularization-based algorithm [19] where BTV is used as the regularization term. The traditional cost function for SR is described as

$$J(X) = \sum_{k=1}^{K} \left\| \mathbf{DHF}^k X - Y^k \right\|_1 + \lambda \sum_{l=-d}^{d} \sum_{m=-d}^{d} \alpha^{|l|+|m|} \left\| X - \mathbf{S}_x^l \mathbf{S}_y^m X \right\|_1, \tag{2}$$

where α is the bilateral filter's parameter which is always pre-determined manually and λ is the regularization parameter.

In the document images, motion can be approximated as pure translation or even it can be adjusted in the system to be pure translation. Then both blur operator and motion operator can commute (i.e. $\mathbf{HF}^k = \mathbf{F}^k\mathbf{H}$). Therefore, the cost function, described in (2), can be modified as

$$J(X) = \sum_{k=1}^{K} \left\| \mathbf{DF}^k\mathbf{H}X - Y^k \right\|_1 + \lambda \sum_{l=-d}^{d} \sum_{m=-d}^{d} \alpha^{|l|+|m|} \left\| X - \mathbf{S}_x^l\mathbf{S}_y^m X \right\|_1 . \quad (3)$$

Assume that $Z = \mathbf{H}X$ is the blurred version of the HR image. The optimization problem described in (3) can be separated into two sub-problems, that is, fusion and restoration:

$$J_1(Z) = \sum_{k=1}^{K} \left\| \mathbf{DF}^k Z - Y^k \right\|_1 , \quad (4)$$

and

$$J_2(X) = \left\| \mathbf{A}(\mathbf{H}X - \hat{Z}) \right\|_1 + \lambda \sum_{l=-d}^{d} \sum_{m=-d}^{d} \alpha^{|l|+|m|} \left\| X - \mathbf{S}_x^l\mathbf{S}_y^m X \right\|_1 , \quad (5)$$

where $\mathbf{A}$ is a diagonal matrix that contains the square root of the number of LR pixels corresponding each pixel position, $\mathbf{S}_x^l$ and $\mathbf{S}_y^m$ are the shifting operators in x and y by l and m respectively, and $((2d+1)\times(2d+1))$ is the support size on which the smoothness is done. The optimum solution ($\hat{Z}$) that minimizes (4) is found to be the median of the up-sampled and warped LR images [19], i.e.

$$\hat{Z}(i) = \underset{k}{\text{median}}(\mathbf{F}^{k^T}\mathbf{D}^T Y^k)(i). \quad (6)$$

This solution implies a fixed mapping function between the LR images and blurred version of the HR image. Also, the reconstructed HR image by using (5) is highly depending on the choice of α and d.

4 Learning-Based Fusion (LBF)

In this step, the LR images are fused to get a blurred version of the HR image, where the mapping function between the LR images and blurred HR image is learned. Assume that the required resolution enhancement factor is r, then, r^2 sub-grids need to be filled in the HR grid. Because of the assumption of pure translational motion, all pixel positions on each sub-grid have the same number of LR pixels to be fused.

Therefore, it is reasonable to update the weights of each sub-grid independently. The mapping function is defined by the weighting parameters W as follows

$$\begin{aligned} \hat{Z}(i) &= \sum_{k=1}^{\mathcal{K}_j} W_j(k)(\mathbf{F}^{k^T}\mathbf{D}^T Y^k)(i), \;\; i \in \mathcal{G}_j, \\ \text{subject to } &\sum_{k=1}^{\mathcal{K}_j} W_j(k) = 1, \end{aligned} \tag{7}$$

where W_j is a $(\mathcal{K}_j \times 1)$ vector that weights the contribution of the LR pixels that aligned to the sub-grid $\mathcal{G}_j$ and $\mathcal{K}_j$ is the number of LR images aligned to that sub-grid.

The mapping parameters are learned so as to minimize the L_1-norm of the difference between the observed LR images and the simulated LR images, where the simulated LR images are generated using the observation model in (1) as follows:

$$W_j = \arg\min_{W_j} \left\{ \sum_{k=1}^{\mathcal{K}_j} \|\mathbf{D}\mathbf{F}^k \hat{Z} - Y^k\|_1 \right\}. \tag{8}$$

Therefore, the updating rule for parameters W_j are defined as:

$$W_j^{n+1}(i) = W_j^n(i) - \mu_{w_j} Y^i(n) \sum_{k=1}^{\mathcal{K}_j} \text{sign}\left((\mathbf{D}\mathbf{F}^k \hat{Z})(n) - Y^k(n) \right), \tag{9}$$

where n is the sample's index of the sub-grid. The mapping parameters are updated at each pixel position of the HR grid. The updating process of the mapping parameters is described as follows; starting at position $(0, 0)$ of the sub-grid $\mathcal{G}_j$, the mapping parameters are initialized as $W_j^0 = [\frac{1}{\mathcal{K}_j} \frac{1}{\mathcal{K}_j} \cdots \frac{1}{\mathcal{K}_j}]$. Then the parameters are updated progressively according to a sequential scanning pattern for each sub-grid. For its stability and fast convergence, the adaptive step size of the generalized NLMS algorithm [22] for L_1-norm case is used to choose the step size of learning parameters. The step size of the mapping function's parameters is chosen as

$$\mu_{w_j} = \frac{\gamma}{c + \|R_{w_j}\|_1}, \tag{10}$$

where $R_{w_j} = [(\mathbf{F}^{1^T}\mathbf{D}^T Y^1)(n), \ldots, (\mathbf{F}^{\mathcal{K}_j{}^T}\mathbf{D}^T Y^{\mathcal{K}_j})(n)]^T$, $c > 0$.

Indeed, the updating rule described in (9) depends on the scanning pattern. However, the scanning pattern has small effect on the fusion step. The reason for this small effect can be explained as follows. Document images are pseudo-binary images which implies that in case of noiseless images (registration error is the only source of error), images' pixels have one of two values (0 or 255). Then the effect

of scanning pattern may happen at transition from 0 to 255 or from 255 to 0. In the first case, pixels that have zero value are not affecting the weights (weights are not updated as shown in (9)). Also in the second transition zero pixels are not affected by the update at pixels that have values equal 255. In case of noisy images, the noise affect the zero values and then affect the updating process, however this effect depends on the amount of contaminating noise, as will be shown later by simulation.

5 Directional Total Variation (DTV) and Adaptation

Instead of assuming smoothness in all directions, by using Tikhonov regularization, or using constant values for α and λ as in (5), we propose to use a general regularization function consisting of a linear combination of the smoothness in different directions with adaptive weight ($\Gamma_{l,m}(X)$). To do that, we propose to replace $\lambda\alpha^{|l|+|m|}$ by a generalized parameter, $\Gamma_{l,m}(X) \geq 0$, which we call the *DTV parameter* in direction (l, m). We want to emphasis that $\Gamma_{l,m}$ is not only a generalization of the conventional non-directional parameter, but also is a function of X. We will discuss how to define this function and how to adapt this parameter later. In fact this regularization function can be viewed as a smoothness in different direction with different regularization parameter.

5.1 Adaptive DTV-Based Restoration

The cost function of the proposed restoration algorithm with DR-based regularization is described as

$$J_3(\Gamma(X), X) = \left\| \mathbf{H}X - \hat{Z} \right\|_1 + \sum_{l=-d}^{d} \sum_{m=-d}^{d} \Gamma_{l,m}(X) \left\| \mathbf{C}_{l,m} X \right\|_1, \tag{11}$$

where $\|\cdot\|_1$ is the L_1-norm, $\Gamma_{l,m}(X)$ is the adaptive weight for direction (l, m) to be learned based on the partially restored image and $\mathbf{C}_{l,m}$ is a high pass operator that works in the direction (l, m).

5.2 Iterative Minimization

The cost function described in (11) is optimized by using steepest descent for restored image estimation and iterative approximation for DTV parameters computation. The coefficients of the DTV function ($\Gamma_{l,m}(X)$) are computed for each direction so as to preserve the convexity of the function $J_3(\Gamma(X), X)$. Assuming $\Gamma_{l,m}(X)$ as a linear function of $J_{l,m}(\Gamma_{l,m}(X), X)$ as

$$\Gamma_{l,m}(X) = \zeta_{l,m} J_{l,m}(\Gamma_{l,m}(X), X), \tag{12}$$

where

$$J_{l,m}(\Gamma_{l,m}(X), X) = \|\mathbf{H}X - Z\|_1 + \Gamma_{l,m}(X)\,\|\mathbf{C}_{l,m}X\|_1\,. \tag{13}$$

Then

$$\Gamma_{l,m}(X) = \frac{\|\mathbf{H}X - Z\|_1}{\frac{1}{\zeta_{l,m}} - \|\mathbf{C}_{l,m}X\|_1}. \tag{14}$$

Therefore, the restored image is updated as

$$X^{(n+1)} = X^{(n)} - \beta\nabla_x J_3(\Gamma(X^{(n)}), X^{(n)}), \tag{15}$$

where $\nabla_x J_3(\Gamma(X), X)$ is computed as follows:

$$\begin{aligned}\nabla_x J_3(\Gamma(X), X) = \mathbf{H}^T\text{sign}\,(\mathbf{H}X - Z) + \sum_{l=-d}^{d}\sum_{m=-d}^{d}\Gamma_{l,m}(X)\mathbf{C}_{l,m}^T\text{sign}\,(\mathbf{C}_{l,m}X)\\ + \sum_{l=-d}^{d}\sum_{m=-d}^{d}\nabla_x\Gamma_{l,m}(X)\|\mathbf{C}_{l,m}X\|_1,\end{aligned} \tag{16}$$

$$\nabla_x\Gamma_{l,m}(X) = \frac{\left(\frac{1}{\zeta_{l,m}} - \|\mathbf{C}_{l,m}X\|_1\right)\mathbf{H}^T\text{sign}\,(\mathbf{H}X - Z) + \|\mathbf{H}X - Z\|_1\mathbf{C}_{l,m}^T\text{sign}\,(\mathbf{C}_{l,m}X)}{\left(\frac{1}{\zeta_{l,m}} - \|\mathbf{C}_{l,m}X\|_1\right)^2}, \tag{17}$$

5.3 Convexity Condition

The sufficient condition for convexity of $J_3(\Gamma(X), X)$ is obtained as follows:

Proposition 1 *The cost function, $J_3(\Gamma(X), X)$, is convex under condition*

$$\frac{1}{\zeta_{l,m}} > \|\boldsymbol{C}_{l,m}X\|_1. \tag{18}$$

Proof is given in the appendix. As a consequence of this condition, $\Gamma_{l,m}$, defined in (14), should be positive value. The first semi-equality can be justified by power preservation of the system, while the second inequality holds when we choose $\mathbf{C}_{l,m}$ as a normalized matrix (i.e. the maximum eigenvalue of $\mathbf{C}_{l,m}^T\mathbf{C}_{l,m}$ equals 1). That is, an appropriate selection of the high-pass operator can guarantee this requirement.

5.4 Convergence analysis

The next step of the analysis is the convergence of the iterative algorithm given in (15). The cost functions consisting of L_1-norm are always solved by utilizing linear

programming [21], however, in this section, the behavior of the proposed iterative minimization, is analysed with simple steepest descent optimization. Assuming that the convexity condition in (18), is satisfied, we can choose $\zeta_{l,m}$ such that

$$\frac{1}{\zeta_{l,m}} = (1+\tau)\|\hat{Z}\|_1 > \|\mathbf{C}_{l,m}X\|_1, \quad \tau > 0, \tag{19}$$

under assumption that

$$\|\hat{Z}\|_1 \simeq \|X\|_1 \geq \|\mathbf{C}_{l,m}X\|_1.$$

Therefore, the behaviour of the cost function, $J_3(\Gamma(X), X)$, is given as follows:

Proposition 2 *The maximum peak of the oscillation of the cost function defined in (11), can be controlled by the step size β and the choice of $\tau \geq 1$, under assumption that $\|\boldsymbol{H}X - \hat{Z}\|_1 = \|V\|_1 \leq \|\hat{Z}\|_1$.*

Proof rewriting the update equation in (15) for two successive iterations using the gradient in (16), we obtain

$$\begin{aligned} X^{(n+1)} - X^{(n)} = X^{(n)} - X^{(n-1)} - \beta\Big\{ & \mathbf{H}^T\big(\mathrm{sign}(\mathbf{H}X^{(n)} - \hat{Z}) - \mathrm{sign}(\mathbf{H}X^{(n-1)} - \hat{Z})\big) \\ & + \sum_{l=-d}^{d}\sum_{m=-d}^{d} \mathbf{C}_{l,m}^T\big(\Gamma_{l,m}(X^{(n)})\mathrm{sign}(\mathbf{C}_{l,m}X^{(n)}) - \Gamma_{l,m}^{(n-1)}\mathrm{sign}(\mathbf{C}_{l,m}X^{(n-1)})\big) \\ & + \sum_{l=-d}^{d}\sum_{m=-d}^{d} \big(\nabla_X\Gamma_{l,m}(X^{(n)})\|\mathbf{C}_{l,m}X^{(n)}\|_1 - \nabla_X\Gamma_{l,m}^{(n-1)}\|\mathbf{C}_{l,m}X^{(n-1)}\|_1\big)\Big\}. \end{aligned} \tag{20}$$

Utilizing (17) and (19), the term $\nabla_X\Gamma_{l,m}\|\mathbf{C}_{l,m}X\|_1$ can be approximated as

$$\nabla_X\Gamma_{l,m}\|\mathbf{C}_{l,m}X\|_1 \leq \frac{1}{\tau}\mathbf{H}^T\mathrm{sign}(\mathbf{H}X - \hat{Z}) + \frac{1}{\tau^2}\mathbf{C}_{l,m}^T\mathrm{sign}(\mathbf{C}_{l,m}X).$$

Then, (20) can be modified as

$$\begin{aligned} & X^{(n+1)} - X^{(n)} \leq \\ & X^{(n)} - X^{(n-1)} - \beta\Big\{(1+\frac{1}{\tau})\mathbf{H}^T\Big(\mathrm{sign}(\mathbf{H}X^{(n)} - \hat{Z}) - \mathrm{sign}(\mathbf{H}X^{(n-1)} - \hat{Z})\Big) \\ & + \sum_{l=-d}^{d}\sum_{m=-d}^{d}(1+\frac{1}{\tau^2})\mathbf{C}_{l,m}^T\Big(\Gamma_{l,m}(X^{(n)})\mathrm{sign}(\mathbf{C}_{l,m}X^{(n)}) - \Gamma_{l,m}^{(n-1)}\mathrm{sign}(\mathbf{C}_{l,m}X^{(n-1)})\Big)\Big\}. \end{aligned} \tag{21}$$

In this inequality, there are four terms containing sign function which has only three possible values, -1, 0 and 1. Therefore, (21) can be written as

$$\begin{aligned} X^{(n+1)} - X^{(n)} \leq X^{(n)} - X^{(n-1)} + \beta\Big\{ & 2(1+\frac{1}{\tau})\mathbf{H}^T 1_1 \\ & + \sum_{l=-d}^{d}\sum_{m=-d}^{d}(1+\frac{1}{\tau^2})\big(\Gamma_{l,m}(X^{(n)})\mathbf{C}_{l,m}1_2 + \Gamma_{l,m}(X^{(n-1)})\mathbf{C}_{l,m}1_3\big)\Big\}, \end{aligned} \tag{22}$$

where 1_1, 1_2 and 1_3 are vectors that consist of any combination of -1, 0 and 1 which result from the corresponding sign functions, Considering the norm for the both sides in (22) and checking the upper bound using triangular inequality, we obtain

$$\begin{aligned}
&\|X^{(n+1)} - X^{(n)}\| \leq \|X^{(n)} - X^{(n-1)}\| + \beta\Big\{\|2(1+\frac{1}{\tau})\mathbf{H}^T 1_1 \\
&+ (1+\frac{1}{\tau^2})\sum_{l=-d}^{d}\sum_{m=-d}^{d}\big(\Gamma_{l,m}(X^{(n)})\mathbf{C}_{l,m}1_2 + \Gamma_{l,m}(X^{(n-1)})\mathbf{C}_{l,m}1_3\big)\|\Big\}, \\
&\leq \|X^{(n)} - X^{(n-1)}\| + \beta\Big\{2(1+\frac{1}{\tau})\|\mathbf{H}^T 1_1\| \\
&+ \|(1+\frac{1}{\tau^2})\sum_{l=-d}^{d}\sum_{m=-d}^{d}\Gamma_{l,m}(X^{(n)})\mathbf{C}_{l,m}1_2 + \Gamma_{l,m}(X^{(n-1)})\mathbf{C}_{l,m}1_3\|\Big\}, \\
&\leq \|X^{(n)} - X^{(n-1)}\| + \beta\Big\{2(1+\frac{1}{\tau})\|\mathbf{H}^T\|\|1_1\| + (1+\frac{1}{\tau^2}) \\
&\sum_{l=-d}^{d}\sum_{m=-d}^{d}\big(\Gamma_{l,m}(X^{(n)})\|\mathbf{C}_{l,m}\|\|1_2\| + \Gamma_{l,m}(X^{(n-1)})\|\mathbf{C}_{l,m}\|\|1_3\|\big)\Big\}.
\end{aligned} \tag{23}$$

Since $\zeta_{l,m}$ is restricted by the convexity condition to be greater than $\|\mathbf{C}_{l,m}X\|_1$, then using (14) together with (19),

$$\Gamma_{h,k}(X^{(n)}) = \frac{\|V\|_1}{(1+\tau)\|Z\|_1 - \|\mathbf{C}_{h,k}X^{(n)}\|_1} \leq \frac{\|Z\|_1}{\tau\|Z\|_1} \leq \frac{1}{\tau}, \tag{24}$$

in a similar way as in (24) we get that $\Gamma_{l,m}(X^{(n)}) \leq \frac{1}{\tau}$ which means that as $\tau \geq 1$, $\Gamma_{l,m}(X^{(n)}) \leq 1$. Then the last term in (23) is bounded by the maximum value which obtained when the values of both 1_1, 1_2 and 1_3 are all ones. Therefore, the parameters β and τ can control the maximum peak of the oscillation of the cost function in (11) if it will happen.

6 Simulation Results and Discussion

Four document images including different types of data available in real documents such as texts, equations, draws and plots, are tested. The evaluation are done as follows. First, the fusion step is tested by fusing LR images. Second, the restoration step is tested by restoring a degraded images. Finally, the complete SR algorithm including fusion and restoration are tested by super-resolving LR images.

6.1 Experiment Setup

In the restoration experiments, the degraded images are generated from original images as follows. The original images are blurred with a Gaussian operator with

kernel size equals 5×5 and variance equals 1 and then a Gaussian noise with variance equals 5 are added to the blurred images.

In the SR experiments, the LR images sequence of the document images are generated from the available HR images as follows. At first, the HR image is blurred with a Gaussian operator with kernel size equals 5×5 and variance equals 1. Then the blurred images are shifted with random values to generate K randomly shifted versions of the HR image ($K = 30$). Finally, the shifted versions of the HR images are down-sampled with factor equals 4 in each direction and an additive Gaussian noise with SNR equals 30 dB is added.

In the evaluation, the restored images using DTV-based restoration is compared with the best and worst results of BTV-based restoration. Also, for more fair comparison, λ in case of BTV-based restoration is adaptively estimated in a similar way as in [23].

In all experiments, the blurring operator is assumed to be known. The parameters used in the simulations are set as follows. Steepest descent is used for optimization. The maximum number of iterations are 100, $\beta = 5$, and $\gamma = 0.5$. $\zeta(l, m, i)$ is chosen as $\frac{1}{2\|\mathcal{P}_i(\hat{Z})\|_1}$, for all l, m values. In the evaluation, the reconstructed HR images using learning-based fusion + DTV-based restoration is compared with the best and worst results of MSA + BTV-based restoration algorithm.

6.2 Fusion Results

An experiment is performed to demonstrate the efficiency of the learning-based fusion step as follows. For Text image, register LR images using translation model based on Lucas-Kanade algorithm [24], then determine the number of images that lay on the same HR grid. In this experiment, we choose the positions $(4 : 4 : L, 4 : 4 : L)$, where three LR images lay on these positions. So we have three LR pixels in each position to be fused.

Three cases are tested in this experiment, the first is adding heavy Gaussian noise with SNR equals 10 dB, to one of these three images. Simply, taking median over these images, containing outliers, can overcome (reject) these outliers. Also, using learning-based fusion can reject these outliers as indicated by small weights corresponding image that contains outliers ($W_1(3)$) as shown in Fig. 1a. The second case is adding heavy noise to two images of these three images. Simply, taking median over these images may lead to a noisy value. However, using learning-based fusion can adjust the weight for each pixel as shown in Fig. 1b, where the weights $W_1(1)$ and $W_1(3)$, corresponding to heavy noisy images, are small (may be zero in some pixel positions) compared to the weight $W_1(2)$ corresponding to the less noisy image. Finally, adding heavy noise to all the three images, increases the probability to get noisy image by using MSA, however, using learning-based fusion can adaptively adjust the weight for each pixel to get the optimal combination of the noisy images as shown in Fig. 1c.

Table 1: The effect of outliers on the PSNR of the fused images.

Sequence	Fusion algorithm	Case I	Case II	Case III	Case IV
Text	LBF with Hilbert scan	17.5484	17.5447	17.3727	17.1706
	LBF with raster scan	17.5049	17.5001	17.3389	17.1008
	MSA	17.5049	17.5049	17.2935	17.0004
Equation	LBF with Hilbert scan	20.5590	20.5243	20.2359	20.7186
	LBF with raster scan	20.3904	20.3652	20.0232	19.5624
	MSA	20.3847	20.3762	19.8714	19.3180
Plot	LBF with Hilbert scan	17.7482	17.6978	17.4766	17.2164
	LBF with raster scan	17.7461	17.7144	17.5718	17.2757
	MSA	17.7403	17.7228	17.5263	17.1284
Draw	LBF with Hilbert scan	18.9604	18.9538	18.7244	18.4367
	LBF with raster scan	18.9599	18.9555	18.7384	18.4517
	MSA	18.9582	18.9531	18.6509	18.2766

To demonstrate the claim that learning fusing parameters is better than using fixed median function to fuse images, we evaluate the peak-signal-to-noise ratio (PSNR) of the fused images in case of using median function and using learning-based fusion. The PSNR of the fused images for different image sequences are reported in Table 1. Four cases are evaluated, namely, using the test sequences without adding more outliers, where the registration error and additive noise in LR images generation process are the only sources of errors (case I), adding heavy noise (SNR$=10$ dB) to one LR of three images that are aligned to the same positions on HR grid (case II), adding heavy noise to two LR images of three LR images (case III), and adding heavy noise for all the three LR images (IV). From this table, we can see that learning fusion parameters enhance the fusion quality especially in case of heavy noise (outliers) contaminating LR images. It worth to mention that in this experiment, the additive heavy noise (outliers) are added to image(s) which correspond to one position of 16 possible positions in the HR grid (resolution enhancement factor is set to 4). Therefore adding more noise (outliers) to different position can show more improvements of the learning-based fusion compared to median-based fusion.

In addition, Table 1 shows a comparison between the PSNR of the fused images using raster scan and Hilbert pattern scan. From this table, we can see that the scanning pattern has small effect on the fusion step in most cases.

6.3 Restoration Results

The iterative restoration step is evaluated in this section. In this evaluation, the restored images using DTV-based restoration is compared with the best and worst

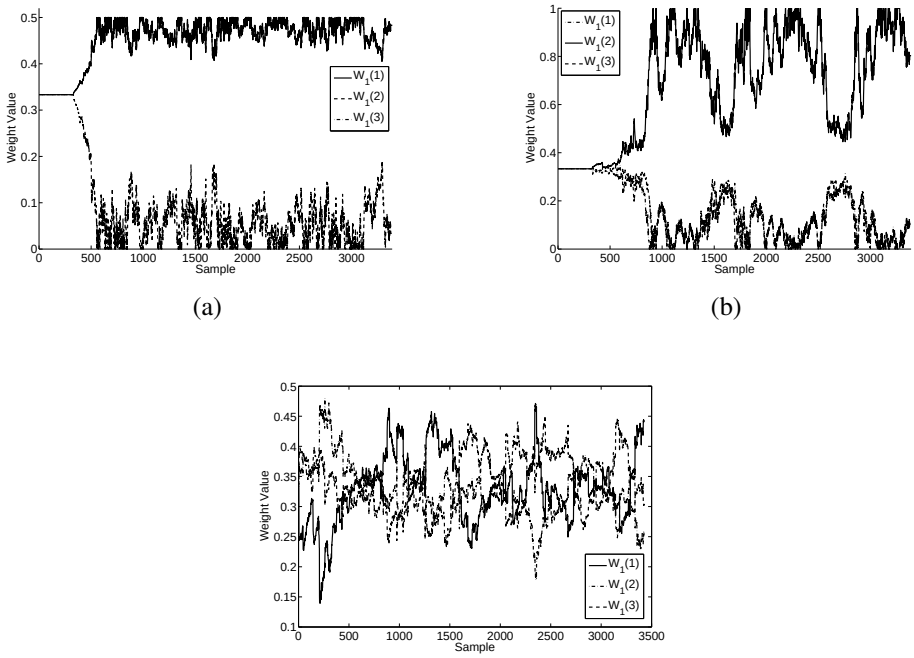


Figure 1: Plot for the fusion parameters in case of: (a) One image of three LR images includes outliers, (b) Two images of three LR images include outliers, (c) All the three images include outliers.

results of BTV-based restoration.

6.3.1 Subjective Results

In Fig. 2, the results of the Text image are shown. A zoomed part for the original image and the degraded image are shown in Figs. 2a and 2b, respectively. The worst restored image, among different selections of α in range $[0.1, 1]$, and the best restored image by using BTV-based restoration are shown in Figs. 2c and 2d, respectively. The restored images by using DTV-based restoration is shown in Fig. 2e. Similarly, the results of Draw image is shown in Fig. 3. From these figures, we can recognize that; BTV-based restoration highly depends on the choice of the bilateral parameter α, using DTV-based restoration can achieve even better quality than BTV-based restoration algorithm with the best choice of α.

6.3.2 Objective Results

To demonstrate the visual results of the proposed restoration algorithm, we use the objective results based on mean absolute error (MAE) evaluation. A comparison

between the MAE values of the restored images by using BTV-based restoration and DTV-based restoration for different document images is shown in Fig. 4. In this figure, the effect of parameters α, and d are evaluated. Parameter α (bilateral parameter) greatly affect the BTV-based algorithm. The parameter d (parameter of the support size of the regularization function) affects both BTV-based and DTV-based restoration algorithms, however, its effect on the BTV-based algorithm is higher.

Finally, to check the convergence of the cost function $J_3(\Gamma(X), X)$ with the simultaneous estimation of $\Gamma(X)$ for each direction, the plot of the cost function versus iteration is tested. Figure 8a shows the plot of $J_3(\Gamma(X), X)$ for four examples. In this figure, we can see that the cost function may oscillate around the minimum value, however, the oscillating value is very small in most cases. This oscillation is due to the update with sign function in the updating rule in (15).

6.4 SR Results

The overall SR algorithm including fusion and restoration steps is evaluated in this section. To test the efficiency of the proposed algorithm, it is compared with MSA-based fusion + BTV-based regularization [19].

6.4.1 Subjective Results

In Fig. 5, the results of the Text sequence are shown. A zoomed part for the original HR and one of the LR images are shown in Figs. 5a and 5b, respectively. The HR image by using MSA with BTV regularization for worst selection (corresponds to the highest MAE value) and best selection of α are shown in Figs. 5c and 5d, respectively. Using MSA can be a good choice in case of less outliers contaminating LR images sequence. However, in case of heavy noisy sequences, MSA can fail to remove all noisy pixels as shown in some noisy pixels in Fig. 5c. On the other hand, the proposed learning-based algorithm can learn the weights of the LR pixels to be included (or even completely excluded) in the HR image. Moreover, a general prior property function is assumed in the regularization term, which is learned simultaneously to overcome the over-smoothness problem and remove contaminating noise, as shown in Fig. 5e.

Another example for super-resolving images that contain draws with different shapes is shown in Fig. 6. A zoomed part for the original HR and the one of the LR images are shown in Figs. 6a and 6b respectively. The HR image by using MSA with BTV regularization is shown in Figs. 6c and 6d. The HR image by using learning-based fusion with DTV-based regularization is shown in Fig. 6e.

6.4.2 Objective Results

To demonstrate the visual results of the SR algorithm, the MAE is used as a quantitative measure for the quality of the resulting HR images. The PSNR is measured for the resulting HR image for different values of α and different values of d for different document image sequences.

Figure 7 shows the MAE of the reconstructed HR image for different document image sequences. In this figure, the MAE of the reconstructed HR image using MSA with BTV regularization is highly depending on the bilateral coefficient (α). On the other hand, learning-based algorithm with DTV adaptively estimate the weight for each direction based on the partially reconstructed HR image.

To check the convergence of the cost function $J_3(\Gamma(X), X)$ with the simultaneous estimation of $\Gamma(X)$ for each direction, the plot of the cost function versus iteration is tested. Figure 8b shows the plot of $J_3(\Gamma(X), X)$ for four examples. In this figure, we can see that the cost function always oscillating around the minimum value. However, the oscillating value is very small in some cases such as Plot, Equation and Draw sequences examples. In other case the oscillating value is recognizable as in Text sequence example. This oscillation is due to the update with sign function in the updating rule in (15).

7 Conclusion

In this paper, we presented a novel learning-based image SR algorithm. The proposed algorithm is divided into two main steps, the first is the non-iterative fusion step, the second is the iterative restoration step. The fusion step is based on learning the mapping function from the LR to HR images. The restoration step is based on adapting the regularization parameter in each direction based on the partially reconstructed image. The convexity condition as well as the convergence conditions are introduced. The proposed algorithm is tested for document image super-resolution application.

8 Open Problem

The work in this paper can be extended for different motion model rather than translational motion. In this case, all the neighboring pixels will be weighted, where both spatial and temporal weights will be used and adapted.

9 Appendix

A real valued functional f on a convex subset $\mathcal{C}$ of a linear vector space is said to be convex if

$$f(az_1 + (1-a)z_2) \leq af(z_1) + (1-a)f(z_2) \tag{25}$$

for all $z_1, z_2 \in \mathcal{C}$ and $0 < a < 1$. Let us rewrite the nonlinear functional $J_3(\Gamma(X), X)$ with only two regularization terms in two different directions as follows

$$J_3(z) = J_3(s, q, r) = s + f_1(J_{0,1}(s, q))q + f_2(J_{1,0}(s, r))r, \tag{26}$$

where

$$\begin{aligned} J_{0,1}(s, q) &= \|\mathbf{H}X - Z\|^1 + \Gamma_{0,1} \|\mathbf{C}_{0,1}X\|^1, \\ J_{1,0}(s, r) &= \|\mathbf{H}X - Z\|^1 + \Gamma_{1,0} \|\mathbf{C}_{1,0}X\|^1, \end{aligned}$$

where $s = \|\mathbf{H}X - Z\|_1$, $q = \|\mathbf{C}_{0,1}X\|_1$ and $r = \|\mathbf{C}_{1,0}X\|_1$. s, q and r are defined within a convex region, i.e. $s, q, r \in \mathcal{C}$. The condition for convexity of nonlinear functional $J_3(\Gamma(X), X)$ translates into the condition

$$J_3(az_1 + (1-a)z_2) \leq aJ_3(z_1) + (1-a)J_3(z_2), \tag{27}$$

or

$$\begin{aligned} &J_3(as_1 + (1-a)s_2, aq_1 + (1-a)q_2, ar_1 + (1-a)r_2) \\ &\leq aJ_3(s_1, q_1, r_1) + (1-a)J_3(s_2, q_2, r_2), \end{aligned} \tag{28}$$

where $z_1 = (s_1, q_1, r_1)^T$ and $z_1 = (s_2, q_2, r_2)^T$, $s_1, s_2, q_1, q_2, r_1, r_2 \in \mathcal{C}$ and $0 < a < 1$. Using (26), the left hand side of (28) can be written as

$$\begin{aligned} &J_3(as_1 + (1-a)s_2, aq_1 + (1-a)q_2, ar_1 + (1-a)r_2) \\ &= as_1 + (1-a)s_2 + \mathcal{L}, \end{aligned} \tag{29}$$

where

$$\begin{aligned} \mathcal{L} = &(aq_1 + (1-a)q_2)f_1(J_{0,1}(as_1 + (1-a)s_2, aq_1 + (1-a)q_2)) \\ &+ (ar_1 + (1-a)r_2)f_2(J_{1,0}(as_1 + (1-a)s_2, ar_1 + (1-a)r_2)). \end{aligned} \tag{30}$$

Similarly, the right hand side of (28) can be rewritten as

$$aJ_3(s_1, q_1, r_1) + (1-a)J_3(s_2, q_2, r_2) = as_1 + (1-a)s_2 + \mathcal{R}, \tag{31}$$

$$\begin{aligned} \mathcal{R} = &aq_1 f_1(J_{0,1}(s_1, q_1) + (1-a)q_2 J_{0,1}(s_2, q_2)) \\ &+ ar_1 f_2(J_{1,0}(s_1, r_1) + (1-a)r_2 J_{1,0}(s_2, r_2). \end{aligned} \tag{32}$$

Therefore, showing the validity of the inequality in (28) requires showing that the right hand side of (32) is greater or equals the right hand side of (30). With the

assumption that the functions f_1 and f_2 are linear and monotonically increasing of $J_{0,1}$ and $J_{1,0}$ respectively, then $\frac{\partial f_1(J_{0,1})}{\partial J_{0,1}} \geq 0$ and $\frac{\partial f_2(J_{1,0})}{\partial J_{1,0}} \geq 0$. With this assumption the prove of the inequality in (28) is simplified to showing that

$$\begin{aligned}&(aq_1 + (1-a)q_2)J_{0,1}(as_1 + (1-a)s_2, aq_1 + (1-a)q_2)\\&+ (ar_1 + (1-a)r_2)J_{1,0}(as_1 + (1-a)s_2, ar_1 + (1-a)r_2)\\&\leq aq_1 J_{0,1}(s_1, q_1) + (1-a)q_2(J_{0,1}(s_2, q_2) + ar_1 J_{1,0}(s_1, r_1) + (1-a)r_2(J_{1,0}(s_2, r_2).\end{aligned} \tag{33}$$

Both $J_{0,1}(s, q)$ and $J_{1,0}(s, r)$ are similar to the function presented in [23] (since one regularization term is used in each of them) which has been proved to be convex with conditions $\frac{\partial f_1(J_{0,1})}{\partial J_{0,1}} < \frac{1}{q}$ and $\frac{\partial f_2(J_{1,0})}{\partial J_{1,0}} < \frac{1}{r}$, respectively. Therefore, the left hand side in the inequality (33) follows the inequality (25), then

$$\begin{aligned}&(aq_1 + (1-a)q_2)J_{0,1}(as_1 + (1-a)s_2, aq_1 + (1-a)q_2)\\&+ (ar_1 + (1-a)r_2)J_{1,0}(as_1 + (1-a)s_2, ar_1 + (1-a)r_2)\\&\leq (aq_1 + (1-a)q_2)[aJ_{0,1}(s_1, q_1) + (1-a)J_{0,1}(s_2, q_2)]\\&+ (ar_1 + (1-a)r_2)[aJ_{1,0}(s_1, r_1) + (1-a)J_{1,0}(s_2, r_2)].\end{aligned} \tag{34}$$

Then the condition in (33) implies

$$\begin{aligned}&aq_1 J_{0,1}(s_1, q_1) + (1-a)q_2 J_{0,1}(s_2, q_2)\\&- aq_1[aJ_{0,1}(s_1, q_1) + (1-a)J_{0,1}(s_2, q_2)]\\&- (1-a)q_2[aJ_{0,1}(s_1, q_1) + (1-a)J_{0,1}(s_2, q_2)]\\&+ ar_1 J_{1,0}(s_1, r_1) + (1-a)r_2 J_{1,0}(s_2, r_2)\\&- ar_1[aJ_{1,0}(s_1, r_1) + (1-a)J_{1,0}(s_2, r_2)]-\\&(1-a)r_2[aJ_{1,0}(s_1, r_1) + (1-a)J_{1,0}(s_2, r_2)]\\&= a(1-a)(q_1 - q_2)[J_{0,1}(s_1, q_1) - J_{0,1}(s_2, q_2)]\\&+ a(1-a)(r_1 - r_2)[J_{1,0}(s_1, r_1) - J_{1,0}(s_2, r_2)] \geq 0.\end{aligned} \tag{35}$$

Since $a(1-a)$ is always positive, the condition in (35) becomes

$$(q_1 - q_2)[J_{0,1}(s_1, q_1) - J_{0,1}(s_2, q_2)] + (r_1 - r_2)[J_{1,0}(s_1, r_1) - J_{1,0}(s_2, r_2)] \geq 0, \tag{36}$$

which is true if both terms $(q_1 - q_2)[J_{0,1}(s_1, q_1) - J_{0,1}(s_2, q_2)] \geq 0$ and $(r_1 - r_2)[J_{1,0}(s_1, r_1) - J_{1,0}(s_2, r_2)] \geq 0$ which implies that $(q_1 - q_2)$ and $[J_{0,1}(s_1, q_1) - J_{0,1}(s_2, q_2)]$ have the same sign and also $(r_1 - r_2)$ and $[J_{1,0}(s_1, r_1) - J_{1,0}(s_2, r_2)]$ have the same sign. One case for $J_{0,1}(s, q)$ to convex is when both $J_{0,1}(\cdot, q)$ and $J_{1,0}(\cdot, r)$ are monotonically increasing functions of q, r regardless of the other axes s, i.e. $\frac{\partial J_{0,1}}{\partial q} \geq 0$ and $\frac{\partial J_{1,0}}{\partial r} \geq 0$. Since $J_{0,1} = s + f_1(J_{0,1})q$ and $J_{1,0} = s + f_2(J_{1,0})r$

$$\frac{\partial J_{0,1}}{\partial q} = \frac{\partial f_1}{\partial J_{0,1}} \frac{\partial J_{0,1}}{\partial q} q + f_1(J_{0,1}), \tag{37}$$

or

$$\frac{\partial J_{0,1}}{\partial q} = \frac{f_1(J_{0,1})}{1 - q\frac{\partial f_1}{\partial J_{0,1}}} \geq 0. \tag{38}$$

Since $f_1(J_{0,1}) \geq 0$, (38) results in

$$\frac{\partial f_1(J_{0,1})}{\partial J_{0,1}} < \frac{1}{q} = \frac{1}{\|\mathbf{C}_{0,1}X\|_1}. \tag{39}$$

And similarly

$$\frac{\partial f_2(J_{1,0})}{\partial J_{1,0}} < \frac{1}{r} = \frac{1}{\|\mathbf{C}_{1,0}X\|_1}. \tag{40}$$

Generalizing this result to multiple regularization terms rather than two terms, the general condition for convexity for the regularization term in direction (l, m) can be described as

$$\frac{\partial f_{l,m}(J_{l,m})}{\partial J_{l,m}} < \frac{1}{q_{l,m}} = \frac{1}{\|\mathbf{C}_{l,m}X\|_1}. \tag{41}$$

It worth to mention that this condition includes the condition of convexity of both $J_{0,1}(s, q)$ and $J_{1,0}(s, r)$.

References

[1] J. Banerjee and C. V. Jawahar, "Super-Resolution of text images using edge-directed tangent field," in Proc. Int. Workshop Document Analysis Systems, September 2008, pp. 76-83.

[2] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," IEEE Signal Process. Magazine, vol. 20, no. 3, pp. 21-36, May 2003.

[3] S.P. Kim and W.Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," IEEE Trans. Image Process., vol. 2, pp. 534-539, October 1993.

[4] M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion and transparency," Journal of Visual Comm. Image Rep. vol.4, pp. 324-335, December 1993.

[5] H. Ur and D. Gross, "Improved resolution from subpixel shifted pictures," CVGIP: Graphical Models Image Process., vol. 54, no. 2, pp. 181-186, March 1992.

[6] B.C. Tom and A.K. Katsaggelos, "Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images," in Proc. Int. Conf. Image Process., October 1995, pp. 2539-2543.

[7] R.C. Hardie, K.J. Barnard, and E.E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," IEEE Trans. Image Process., vol. 6, pp. 1621-1633, December 1997.

[8] R.R. Schulz and R.L. Stevenson, "Extraction of high-resolution frames from video sequences," IEEE Trans. Image Process., vol. 5, no. 6, pp. 996-1011, June 1996.

[9] H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," Journal of Opt. Society of America, vol. 6, pp. 1715-1726, November 1989.

[10] A.J. Patti, M.I. Sezan, and A.M. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," IEEE Trans. Image Process., vol. 6, pp. 1064-1076, August 1997.

[11] A. Papoulis, "Generalized sampling theorem," IEEE Trans. Circuits Syst., vol. 24, pp. 652-654, November 1977.

[12] M. Trimeche, R. C. Bilcu, and J. Yrjanainen, "Adaptive outlier rejection in image super-resolution," EURASIP Journal Applied Signal Process., vol. 2006, Article ID 38052, 2006.

[13] W.T. Freeman, E.C. Pasztor, and O.T. Carmichael, "Learning low-level vision," Int. Journal Computer Vision, vol. 40, no. 1, pp. 24-57, October 2000.

[14] D. Capel and A. Zisserman, "Super-resolution from multiple views using learnt image models," In Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, December 2001, pp. 627-634.

[15] D. Datsenko and M. Elad, "Example-based single document image super-resolution: a global MAP approach with outlier rejection," Multidimensional Systems and Signal Process., vol. 18 , no. 2-3, pp. 103-121, September 2007.

[16] K. In Kim, M. O. Franz, and B. Scholkopf, "Kernel hebbian algorithm for single-frame super-resolution," in Proc. Statistical Learning in Computer Vision, May 2004, pp. 135-149.

[17] Y. Lu, M. Inamura, and M. del C. Valdes, "Super-resolution of the undersampled and subpixel shifted image sequence by a neural network," Int. Journal Imaging Systems and Technology, vol. 14, no. 1, pp. 8-15, June 2004.

[18] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure transnational motion and common space invariant blur," IEEE Trans. Image Process., vol. 10, no. 8, pp. 1187-1193, August 2001.

[19] S. Farsiu, D. Robinson, M. Elad, P. Milanfar, "Robust shift-and-add approach to super-resolution," in Proc. SPIE Conf. Applications of Digital Signal and Image Processing, August 2003, pp. 121-130.

[20] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multi-frame super-resolution," IEEE Trans. Image Processing, vol. 13, no. 10, pp. 1327-1344 , October 2004.

[21] H. Fu, M. K. Ng, M. Nikolova, J. L. Barlow, "Efficient minimization methods of mixed l2-l1 and l1-l1 norms for image restoration," in Proc. SIAM Scientific Computing, vol. 27, no. 6, pp. 1881-1902, 2006.

[22] S. C. Douglas, "A family of normalized LMS algorithms," IEEE Signal Process. Letters, vol. 1, no. 3, pp. 49-51, March 1994.

[23] M. G. Kang and A. K. Katsaggelos, "General choice of the regularization functional in regularized image restoration," IEEE Trans. Image Process., vol. 4, no. 5, pp. 594-602, May 1995.

[24] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in Proc. Int. Joint Conf. Artificial Intelligence, April 1981, pp. 674-679.

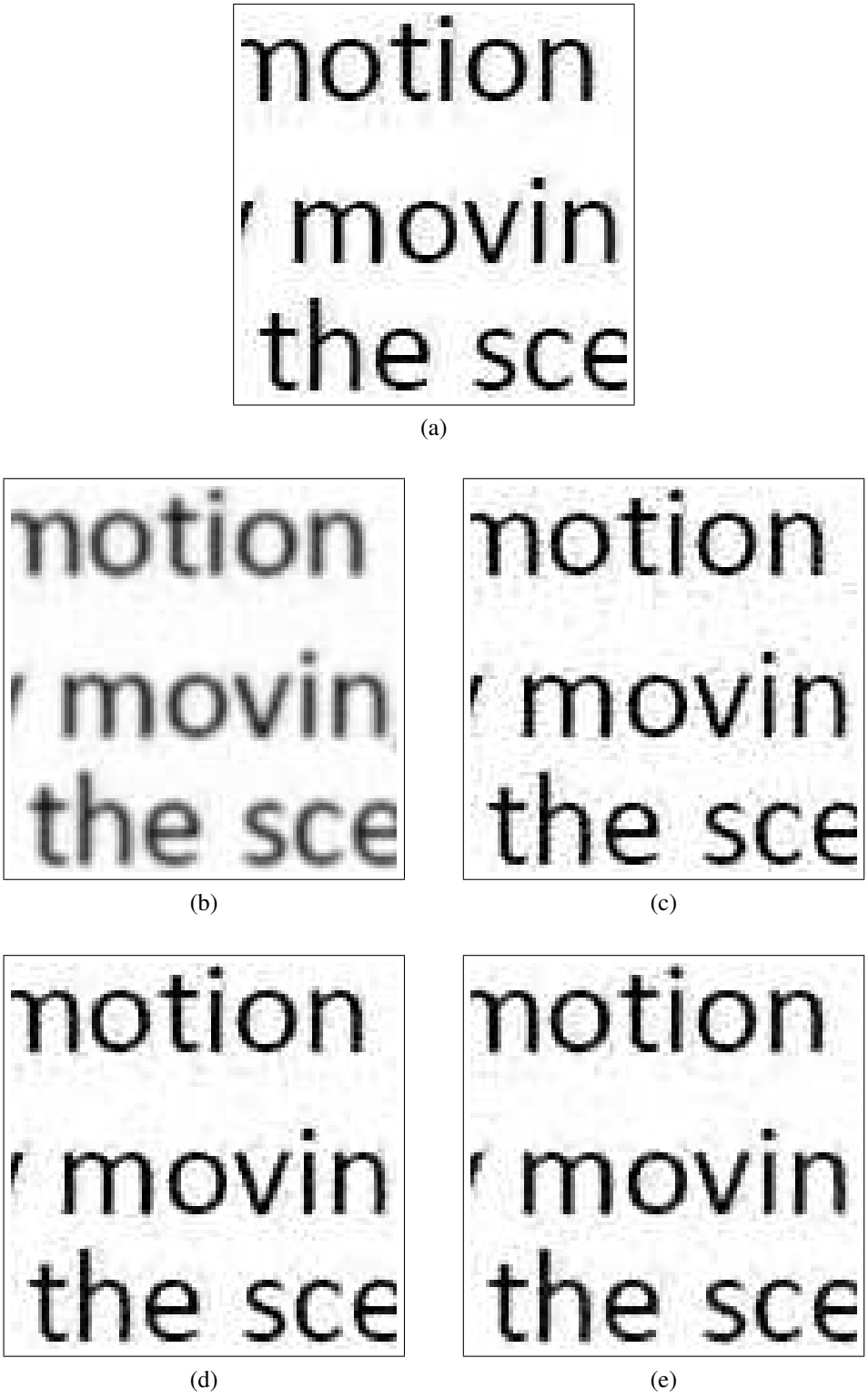
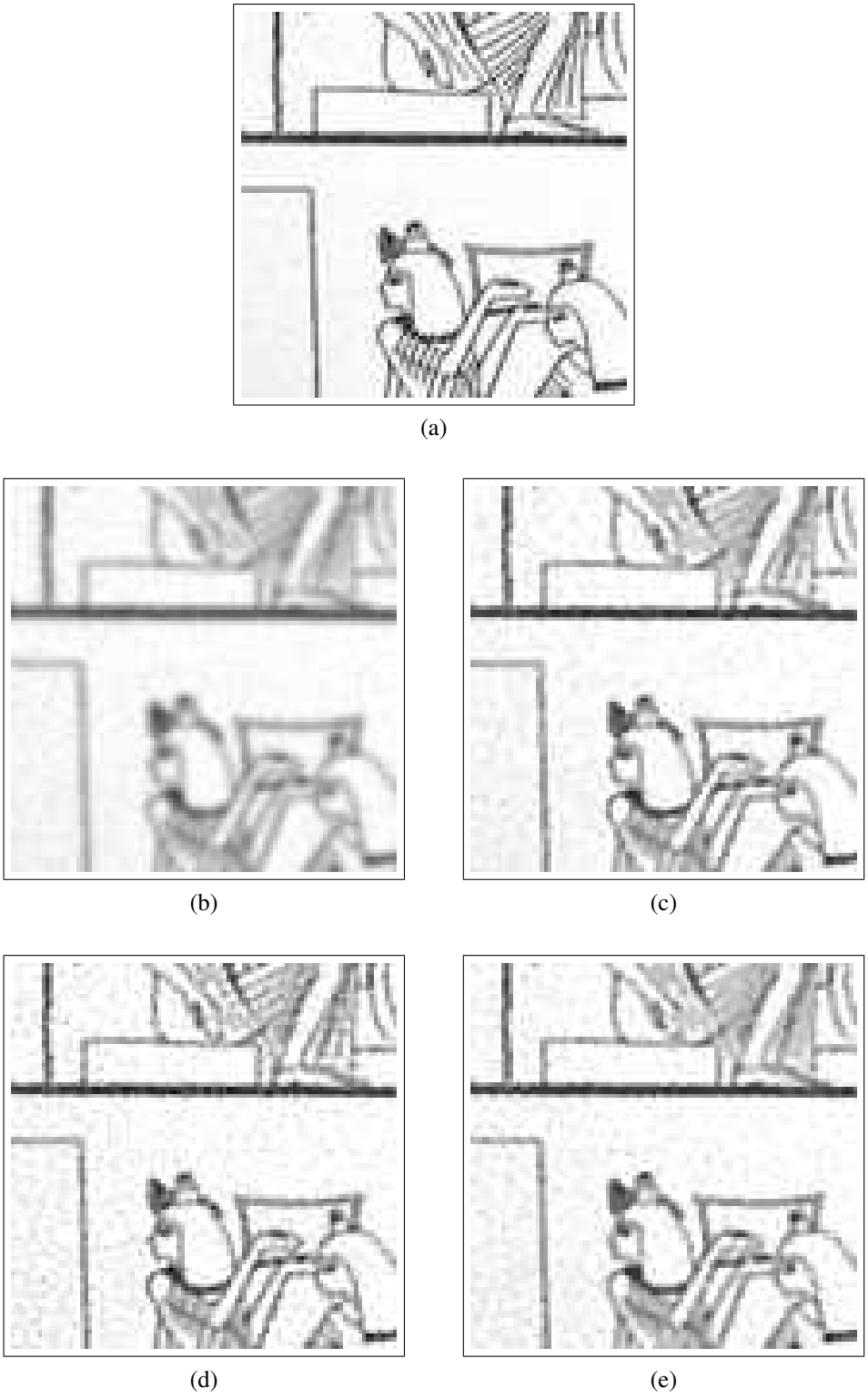


Figure 2: Restoration of Text image: (a) Original image, (b) Degraded image, restored image using; (c) BTV-based restoration [20] corresponding maximum MAE value, (d) BTV-based restoration [20] corresponding minimum MAE value, (e) patch-based DR restoration algorithms using $p = 1$.



(a) (b) (c) (d) (e)

Figure 3: Restoration of Draw image: (a) Original image, (b) Degraded image, restored image using; (c) BTV-based restoration [20] corresponding maximum MAE value, (d) BTV-based restoration [20] corresponding minimum MAE value, (e) patch-based DR restoration algorithms using $p = 1$.

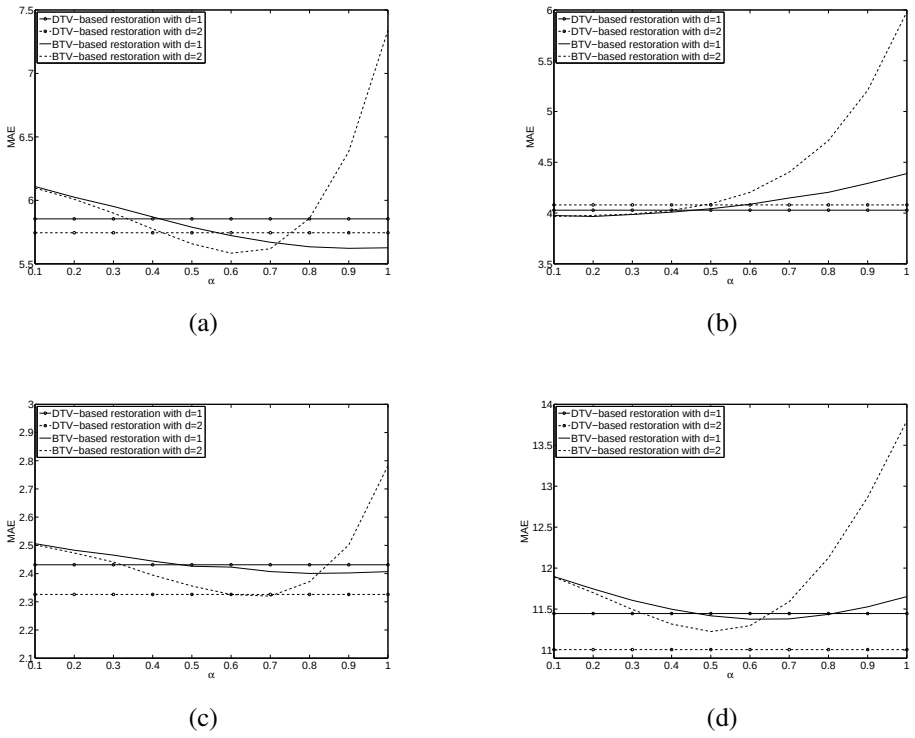
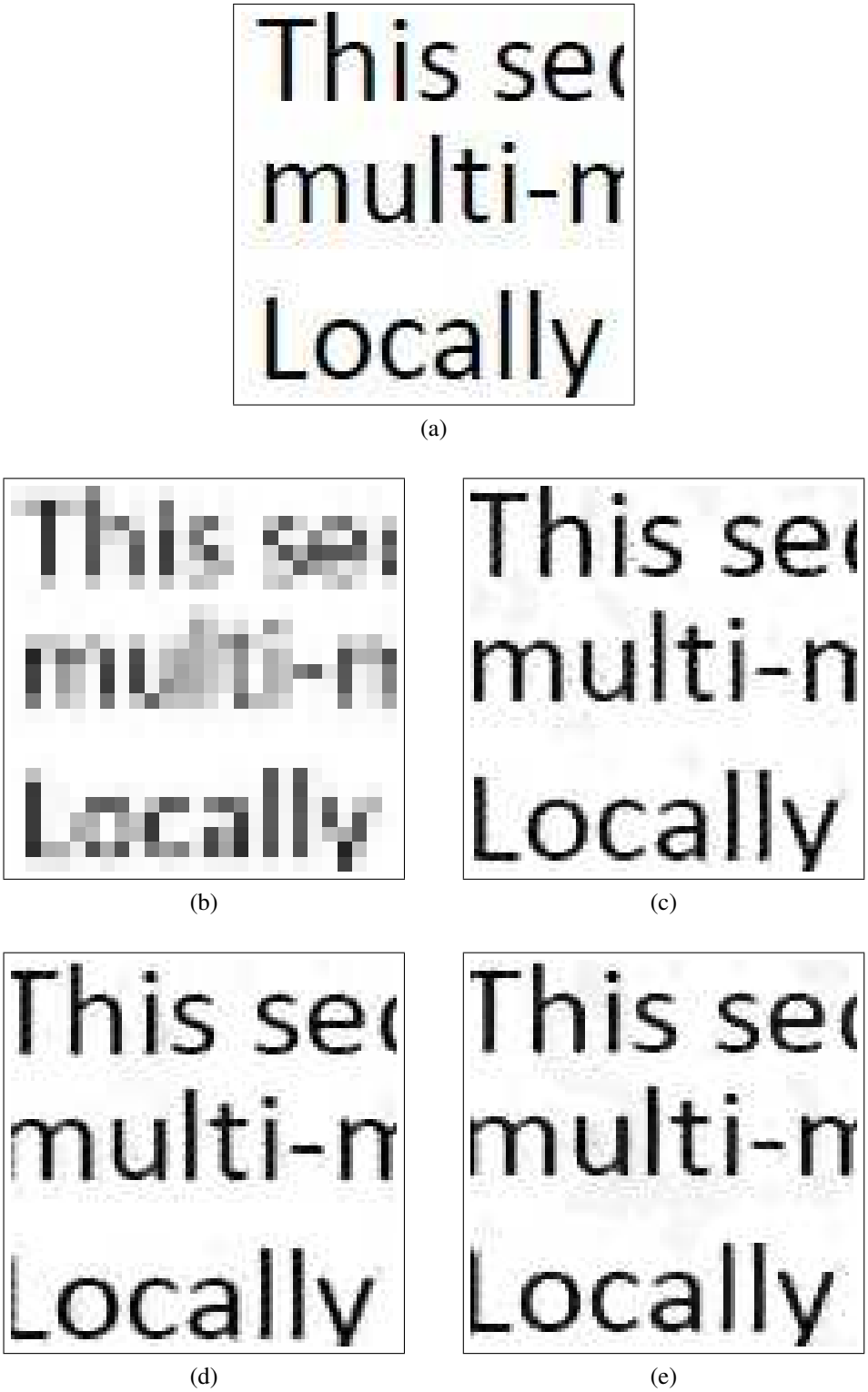


Figure 4: Comparison between the MAE of the restored images by using BTV-based restoration and DTV-based restoration for: (a) Text image, (b) Plot image, (c) Equation image, and (d) Draw image.



(a)

(b)

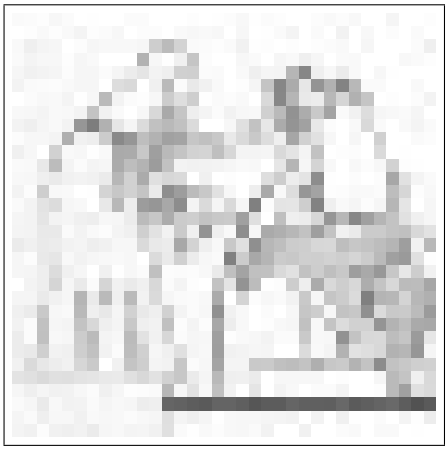
(c)

(d)

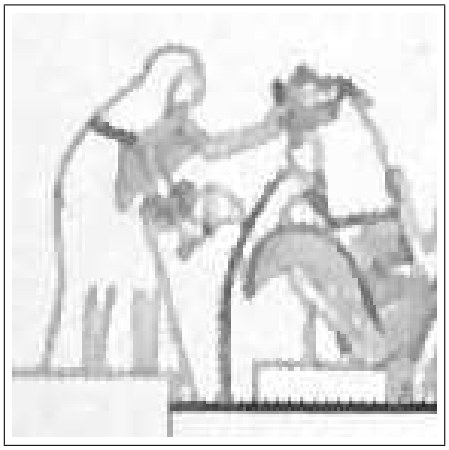
(e)

Figure 5: SR of Text image sequence: (a) Original HR image, (b) LR image, Super-resolved image using; (c) MSA + BTV restoration [19] with worst selection of α, (d) MSA + BTV restoration [19] with best selection of α, (e) Learning-based fusion + patch-based DTV restoration algorithms.

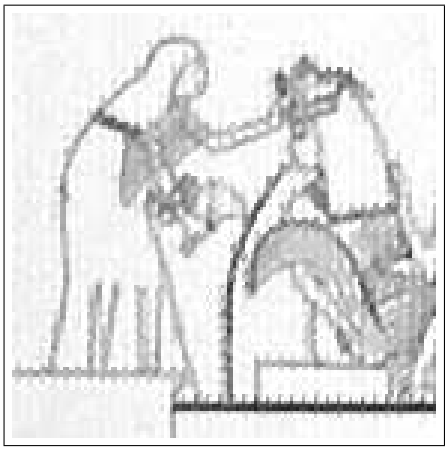
(a)



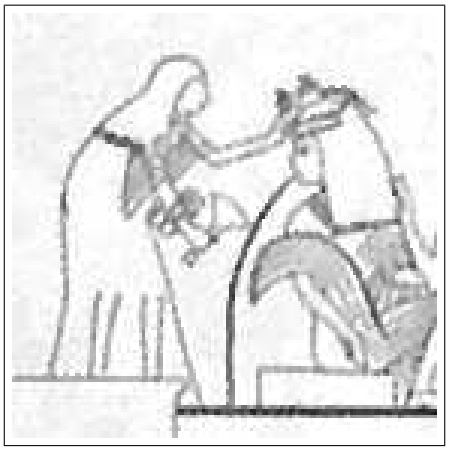
(b)



(c)



(d)



(e)

Figure 6: SR of Draw image sequence: (a) Original HR image, (b) LR image, Super-resolved image using; (c) MSA + BTV restoration [19] with worst selection of α, (d) MSA + BTV restoration [19] with best selection of α, (e) Learning-based fusion + patch-based DTV restoration algorithms.

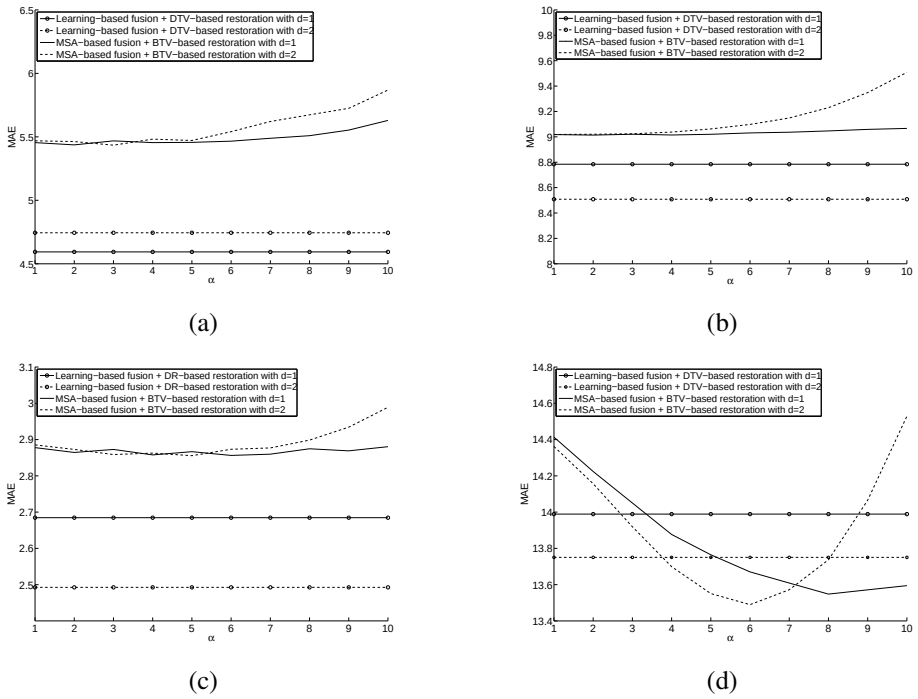


Figure 7: Comparison between the MAE of the reconstructed HR images by using learning-based fusion + DTV-based restoration and MSA-based fusion + BTV-based restoration for: (a) Text image sequence, (b) Plot image sequence, (c) Equation image sequence, and (d) Draw image sequence.

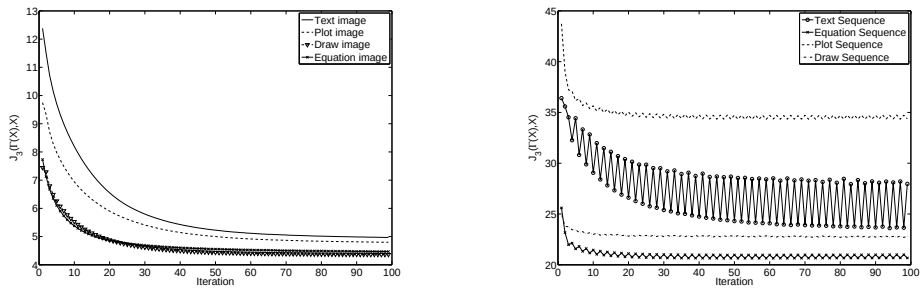


Figure 8: Convergence of function $J_3(\Gamma(X), X)$ in case of: (a) restoration experiment and (b) SR experiment.